Application of Geographically Weighted Regression Analysis on World's Gasoline Prices Model

Tuti Purwaningsih, and Devina Gilar Fitri Ayu Sumardi
Statistics Department
Universitas Islam Indonesia
Jl Kaliurang KM14.5, Yogyakarta, Indonesia
tuti.purwaningsih@uii.ac.id

Abstract

This study aims to determine the effect of inflation, disaster risk, per capita GDP and global import export of world gasoline prices in 2017. The research is a kind of explanatory research with quantitative approach. The data used are nominal data with 5 independent variables namely inflation, GDP per capita, disaster risk, export and import in 2016 with the dependent variable of world gasoline price in 2017. The data used as many as 100 countries in the world are used as samples. With the method of spatial regression and weighted regression of geography the researcher wanted to know the influence of the five independent variables to the world gas price in 2017. The results of this study indicate that simultaneously (t test) using GWR analysis on GDP per capita variables and disaster risk significantly influence the world gasoline prices in 2017, while in the variable of inflation, exports, and imports do not have a significant effect on the world gas price in 2017. Based on the results of research, it is expected that decision makers and the authorities as a policy regulator can take the decision to control macroeconomic conditions in the country so that gasoline prices do not go affected and worsening economic conditions in a country. Similarly, existing disaster risk factors can be considered to make alternatives and other policies to save gasoline prices to keep them stable. Thus, the results of this study can also be used to gain attention in maintaining the stability of petrol and macroeconomic prices remain controlled.

Keywords

Global Oil, Gasoline, GWR, Spatial Regression

1. PRELIMINARY

1.1 Background

Fuel oil is a basic necessity in industries around the world, but fuel oil is a non-renewable natural resource. The need for fuel oil both in the field of industry and transportation is increasing every day because these machines require fuel oil. Because of that, the development of the world oil prices experienced many changes.

Oil and price fluctuations have a vital influence on almost all macroeconomic activity, because oil is one of the major energies used both directly and indirectly in producing goods and services. Oil becomes the top energy source in its use to sustain the production process compared to other energy sources, so that fluctuations in oil prices are very sensitive to economic conditions or economic growth in each country. And no one country is not dependent on oil and is able to instantly reduce its consumption due to price increases. The need for energy in the form of oil is very large, which is a problem that not all countries in the world have oil resources. In another sense the spread of the world's oil reserves is not evenly distributed.

World oil prices are difficult to predict because the factors are numerous, not just technical factors such as demand and supply but also non-technical factors such as war, disaster risk, inflation, GDP of each country, exports, imports and so on. Researching the price of oil is very interesting to discuss because the ups and downs of oil prices will also affect the condition of the internal economy of the Indonesian nation.

Based on the background of the above issues the authors are interested to conduct research entitled "Application of Geographically Weighted Regression Analysis on World's Gasoline Prices Model".

1.2. State of the Problems

Based on the above background, then the problems that arise are:

1. Are there any significant factors affecting the price of gasoline in the world?
2. How is the world macroeconomic condition of 2017 seen from world gas prices, inflation, disaster risk, exports, imports and GDP in each country?
3. How the influence of significant variables from the mapping and data visualization

1.3. Scope of Problem

The problem limitation in this research are:

1. The model used in Geodes is spatial (SAR, SEM, Spatial LAG).
2. The GWR test uses the Gaussian model and the fixed-Gaussian kernel.
3. Mapping and visualizing data using Excel, Geoda, and QuantumGIS.

1.4. Research Purposes

The purpose of this study is to answer the problem formulation namely:

1. Knowing the factors that significantly affect the price of gasoline world 2
2. Knowing the macroeconomic conditions of the world in 2017 seen from world oil prices, inflation, disaster risk, exports, imports and GDP in each country.
3. Knowing the influence of significant variables from mapping and data visualization

3. RESEARCH METHODOLOGY

The research used by type of data and its analysis is quantitative research, that is research which its analysis more focus on numerical data which is processed by spatial regression method using GEODA software.

3.1.Data Collection

In this study, we tested 5 secondary data variables obtained from several official websites such as globalpetrolprices, tradingeconomics, and statbureau, which provides free data for statistics analysis, visualization and sharring. The data used in this study is the world data with variables as in Table 2 below:

Table 1. Research Data

No	Variable	Variable Code	Variable Type	Unit	Source
1	Gasoline Prices	Gasoline	Dependen	USD	www.globalpetrolprices.com
2	Inflation	Inflation	Independen	Percent	www.statbureau.org
3	GDP percapita	GDP	Independen	USD	www.investopedia
4	Disaster Risk	Risk	Independen	Percent	en.wikipedia.org
5	Export	Ekspor	Independen	USD	tradingeconomics.com
6	Import	Impor	Independen	USD	tradingeconomics.com

3.2. Method of analysis

The data that have been obtained are then analyzed in order to find the influence of each variable on the price of gasoline world using GWR method. Then mapped out the results for easy understanding using Geoda and QGis which are then interpreted according to the results obtained.

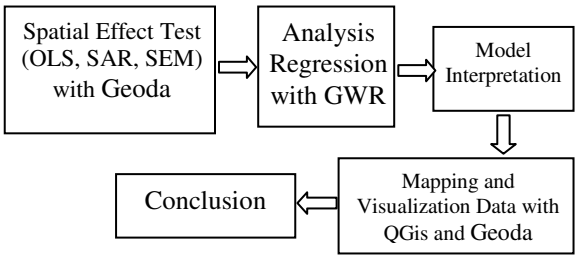

Figure 1. research Flow Diagram

4. RESULTS AND DISCUSSION

4.1. Spatial effect test

Spatial effect testing is performed to determine spatial heterogeneity and spatial dependence. To test spatial heterogeneity was used Breush-Pagan test. From the results of spatial effect testing using Geoda, the results obtained as in table 2 below:

Table 2. Breusch Pagan Test Result

Methods	Breusch Pagan P-Value	Decision
OLS	0,57837 > α	failed reject Ho
SLM	0,45424 > α	failed reject Ho
SEM	0,15719 > α	failed reject Ho

From the results in table 2 above can be concluded that by using the 0.05 significance level of existing data failed reject H0 means regression model contains the existence of homoskedastisitas. Thus, no spatial effect is found in the data, so the influence of the location does not need to be incorporated into the model and the model should no longer have spatial analysis using GWR, but to ensure that the researcher still analyzes using GWR to see which factors affect the price of gasoline in the world.

4.2. Geographically Weighted Regression (GWR)

After doing the analysis of homoschedasticity checking with Geoda, then tested GWR with Gaussian model of the five factors influence the price of gasoline (X) to the dependent variable (Y) the price of gasoline in units of USD obtained the following results:

```
-2 log-likelihood:                  712.757305
Classic AIC:                        726.757305
AICc:                               727.974696
BIC/MDL:                            744.993496
CV:                             4162683.487172
R square:                            0.999982
Adjusted R square:                   0.999981

Variable              Estimate    Standard Error      t(Est/SE)
-------------------- ---------- --------------- ---------------
Intercept          11600.476740      966.527019       12.002227
inflation            -19.088486        8.782243       -2.173532
GDP                    0.125419        0.030130        4.162593
risk                -211.880586       67.576342       -3.135426
ekspor                -0.723626        4.922837       -0.146994
impor                  7.128802        5.961291        1.195849
```

Figure 2. Output GWR 1

From the output in Figure 2 above, it is known that the R square value obtained is 0.999982 indicating that the contribution percentage of independent variable (gasoline price) to the dependent variable (factor affecting the price of gasoline) is 99.9%. While the rest of 0.1% influenced by other variables that are not

included or not discussed in this study. Thus this shows that there is a very strong relationship of factors influencing the price of gasoline against gasoline prices.

Then from the output can also be known which factors that influence significantly and don't have a significant effect. By doing the hypothesis test as follows:

i.Hypothesis:

H0 : The variable did not significantly affect the price of gasoline

H0 : The variable significantly affect the price of gasoline

ii.Level of Significance :

$\alpha = 1\% = 0{,}01$; t table = 1,645

iii.Critical Areas :

t(est/es) < -1,645 ; t(est/es) > 1,645

iv.Decision

Table 3. Test Of Statistics GWR 1

Variabel	T_{hit}	Keputusan
Inflasi	0,117281> t_{tab}	Failed to reject Ho
Risk	4,282378> t_{tab}	Reject Ho
PDB	2223,120> t_{tab}	Reject Ho
Ekspor	0,270589> t_{tab}	Failed to reject Ho
Impor	-0,02528> t_{tab}	Failed to reject Ho

v.Conclusion :

Using the 0.01 significance level, there are 2 significant variables, namely GDP (GDP), and (risk) of disaster risk and 1 intercept variable used as the estimator.

From the first GWR analysis result, there are 3 unambiguous variables, namely inflation, export and import. Inflationary variables did not have a significant effect because inflation rate in each country was included in the category of mild inflation because the rate is less than 10% per year on average. It can be assumed that the economic situation in each country can be said to be safe. While the export and import variables are unlikely to be influenced by the extent of export and imported commodities where the relation to gasoline prices is very low, this export and import should be more focused on the petroleum sector. Because of this, it is necessary to test the GWR again without including the variables that are not significant, to obtain the following results:

```
R square:                              0.999982
Adjusted R square:                     0.999982

Variable                 Estimate    Standard Error      t(Est/SE)
-------------------- --------------- --------------- ---------------
Intercept                -14.588668        1.191427      -12.244697
GDP                        0.000177        0.000041        4.352850
risk                     205.652925        0.090840     2263.897177
```

Figure 3. Output GWR 2

From the result of output of GWR 2 above, by doing the same hypothesis step as before can be directly drawn conclusion that by using level of signifikansi 0,01 all variables have significant effect to gasoline price (Y), that variable, GDP (GDP) and (risk) disaster risk and 1 intercept variable used as an estimate. From the output GWR can also be known the value of coefficient of determination (R2) that is equal to 0.999982, the number has the meaning that this model has independent variables that can explain / affect the dependent variable of 99.9%, while the rest is

affected / explained by other variables outside of this study. Then to see the spatial effects of these two influential variables can be tested using geographically variability test on GWR. From the test results obtained as follows:

```
 Geographical variability tests of local coefficients
********************************************************************************
 Variable             F                  DOF for F test  DIFF of Criterion
-------------------- ------------------ ---------------- -----------------
Intercept                   11.445228    6.560   76.527       -46.874092
GDP                          3.906454    4.966   76.527        -5.996807
risk                        95.889026    3.163   76.527      -149.379689
-------------------- ------------------ ---------------- -----------------
Note: positive value of diff-Criterion (AICc, AIC, BIC/MDL or CV) suggests no
spatial variability in terms of model selection criteria.
  F test: in case of no spatial variability, the F statistics follows the F
distribution of DOF for F test.
```

Figure 4. Result of *Geographically variability test* by GWR

From the above results obtained the value of diff of criterion minus (-) on all variables. So it can be said there are spatial effects on the model under test.

a. Model Interpretation

From the GWR analysis results obtained local model of regression taken for example in one country namely Australia as follows:

Tabel 4. Result of Australian Local Model

Variabel	Est	SE	t
Intercept	1,119686	3,838161	0,291725
GDP	-2,27051	19,99889	-0,11353
Risk	1,05	1,075533	-0,02553

The above variables are used in the model because it has a significant effect on the price of gasoline in Australia. The interpretation, first seen from the value of odds ratio for GDP is $e^{-2,27051}$= 0.103259, from that figure it can be interpreted that the Australian state's chances of having a high gasoline price of 0.103259 times the Australian state's opportunity to have low gasoline prices if the percentage of GDP an increase of 1% from before.

As for the Risk coefficient, seen from the odds ratio is $e^{-1,05}$ = 2,857651, from that number can be interpreted that the opportunity of Australia state have high gasoline price equal to 2,857651 times Australia country opportunity to have low gasoline price if percentage of Risk increase 1 % from the previous.

Interpretations for other countries depend on the local model. Judging from which variable is significant, then to know the effect seen from the odds ratio ratio first. For more details as a whole it can be seen through the mapping of influential variables.

b. Data Visualization

Data visualization includes region mapping, and presents plot / graph data. This is done to see clearly which countries have disaster risk levels and GDP in each country which can then be compared with the price of gasoline in the country to see its influence. Here are the mapping results:

i. Level of Disaster Risk in 2016

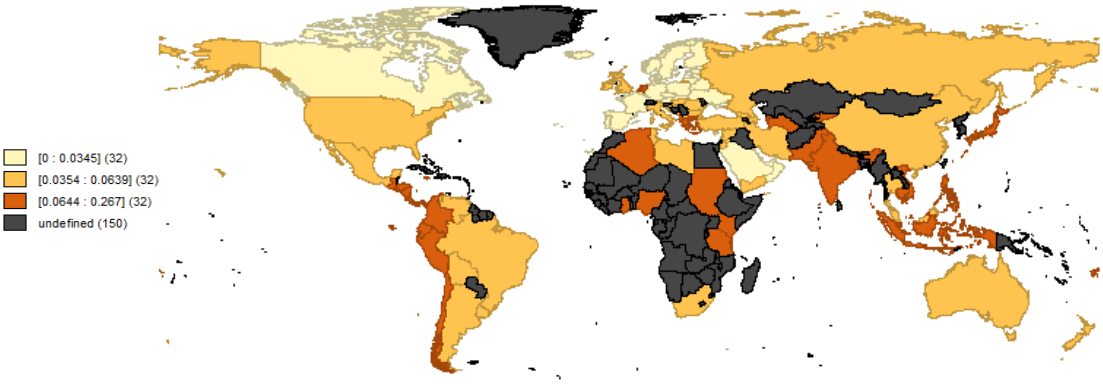

Figure 5. Quantile 3 Mapping of Risk Variable

Based on Figure 5 above, it is known that solid color illustrates the high level of disaster risk, then orange color illustrates the level of risk of moderate disaster, and young orange low disaster risk level. High levels of disaster risk are found in many parts of Asia, south america and africa. While that is almost in all continents, except in the North American dominant has a low level of disaster risk. The order of countries that have disaster risk from highest to low is presented in the following graph,

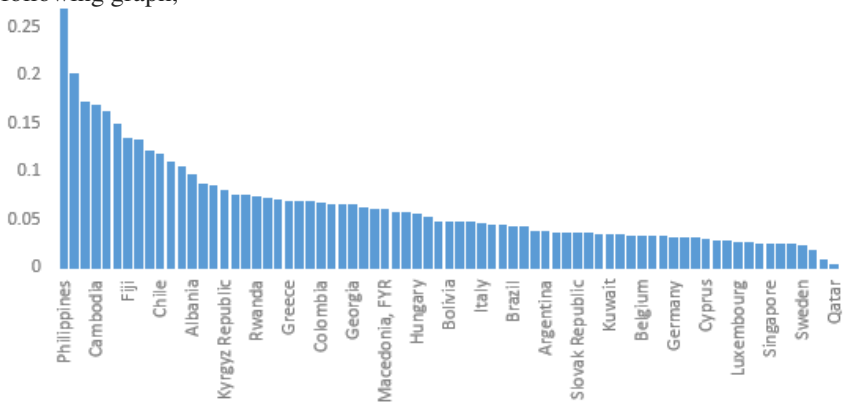

Figure 6. Graph of International Disaster Risk in 2016

From the graph it is known that the country with the highest disaster risk level is Philippines and the lowest in Qatar.

ii. Level of GDP percapita in 2016

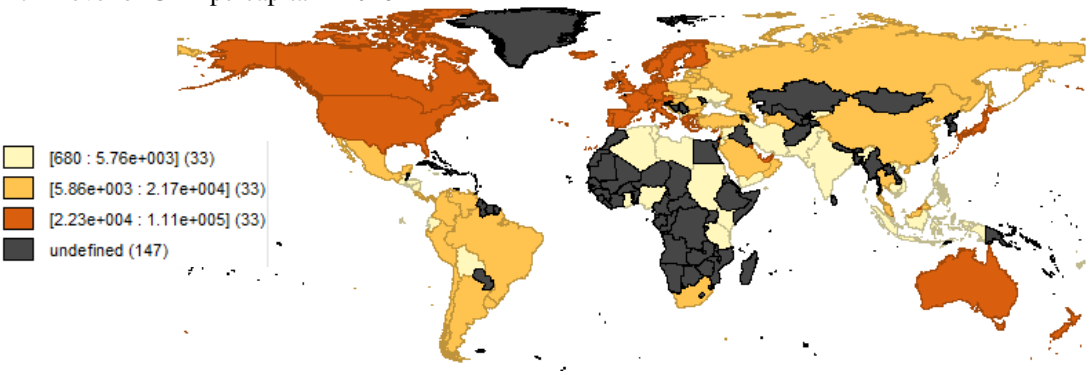

Figure 7. Mapping Quantile 3 of Variables GDP percapita

Based on Figure 7 above, it is known that solid color represents a high level of GDP, then orange color illustrates medium GDP levels, and young orange GDP levels are low. High levels of GDP are found in the continent of Australia, North America and parts of Europe. While that is dominant in the south american and asia region, except in Africa and northern part of africa have low GDP level. The order of countries that have GDP from the highest to low is presented in the following graph,

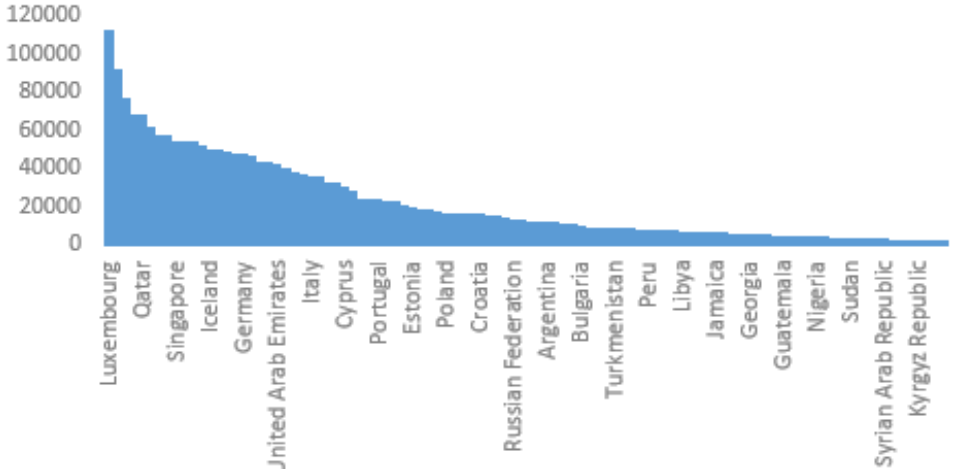

Figure 8. Graph of GDP percapita in the World in 2016

From the graph it I known that the country with the highest risk level is Luxemburg and lowest at Kyrgyz Republic.

iii. Level of Gasoline Prices in 2017

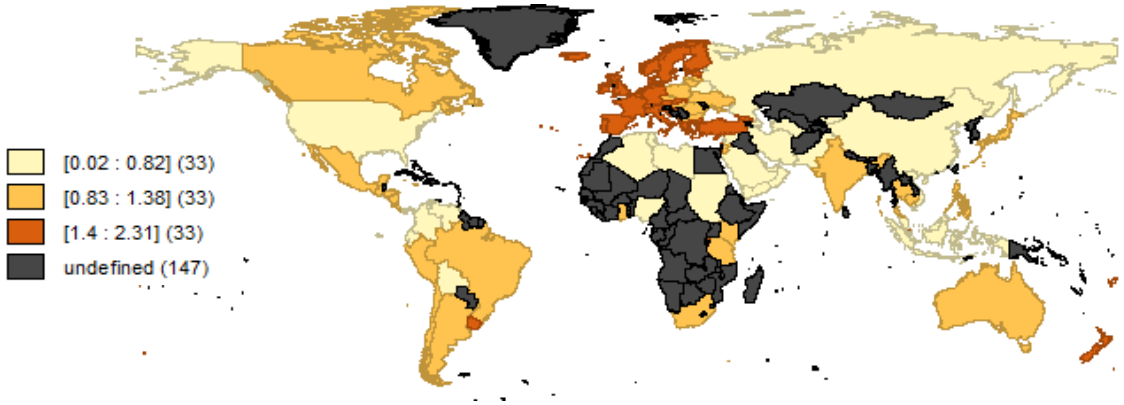

Figure 9. Mapping Quantile 3 of Variables Gasoline Prices in 2017

Based on the above figure 9, it is known that solid color depicts the high gasoline price level, then orange color illustrates the moderate gasoline price level, and the young orange gasoline price level is low. High gasoline prices are dominant in the European continent. While the middle, dominant in the south american region, Australia and some North America. In asia and parts of north america and most of Africa have low gasoline price levels. The order of countries that have the price of gasoline from the highest to low is presented in the following graph,

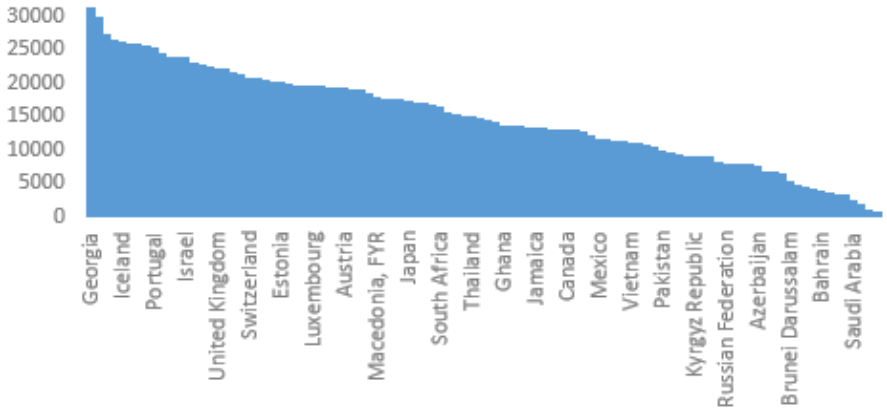

Figure 10. Graph of (IDR) World Gasoline Prices in 2017

From the graph it is known that state with the highest gasoline prices in Georgia and the lowest in Saudi Arabia. Furthermore, to see the comparison graph of the three variables, the researcher uses 3-dimensional graphs on Geoda, following the output :

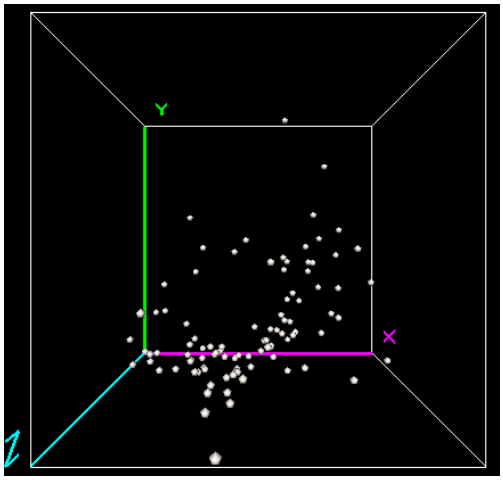

Figure 11. XY Side Graph 3 Dimensions Based on Variable X (Gasoline) Y (GDP) Z (risk)

From 3-dimensional graph above, it can be seen the relationship pattern of gasoline price data (X) with GDP per capita (Y), where the pattern spread following pattern which tend to positive correlation, where the higher the GDP then the price of gasoline tends to be high also. But the correlation value is not strong enough, because masi many spread of data that is not controlled (outlier).

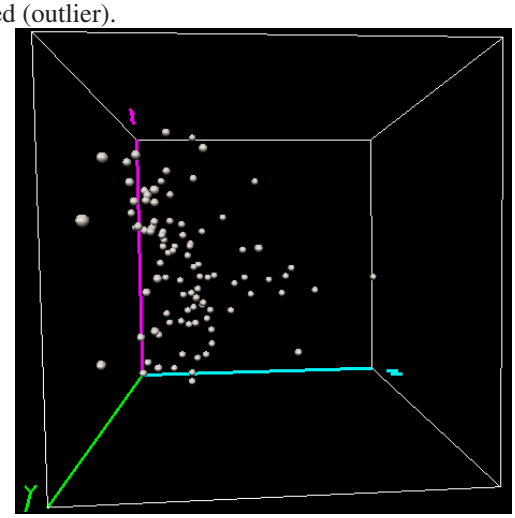

Figure 12. XZ Side Graph 3 Dimensions Based on Variable X(gasoline) Y(GDP) Z(risk)

While the figure 12 shows the variable side of the price of gasoline (X) to disaster risk (Z) which from the graph shows that the data pattern tends not to correlate, because many are assembled flat upwards. However, if seen the spread of data, there is a tenuous pattern that follows the pattern of negative correlation, where the lower the price of gasoline risks higher disaster.

5. CONCLUSIONS

Based on the results of the analysis that has been done in this study, obtained some conclusions to answer the problem formulation as follows:

1. From five independent variables, namely Inflation, disaster risk, GDP, exports, and imports, there are 2 variables that significantly influence the world price of gasoline that is disaster risk variable and GDP per capita. Of the two variables also obtained spatial effects that are evidenced by the results of geographical variability test using GWR.

2. From the results of mapping based on 2 variables that have a significant effect, known macroeconomic distribution of countries in the world. Mecroeconomic conditions tend to be high in the continent of Australia, North America and parts of Europe. While that is dominant in the south american and asia region, except in Africa and northern part of africa have low GDP level.
3. Viewed from the mapping results based on GDP and gasoline prices, the tendency of the high level is the same low, ie the dominant high is found in the European continent, while the moderate, the dominant in southern america, Australia and parts of North America. In Asia and most of Africa have low gasoline prices and low GDP.

REFERENCES

Cliff, A., dan J.K. Ord. 1973. Spatial Autocorrelation. London: Pion.

GlobalPetrolPrices (2017). Gasoline Prices. http://www.globalpetrolprices.com. Accessed on January 16th 2018 at 11.14 am.

Lee J.and Wong D.W.S. 2001, Statistic for Spatial Data, John Wiley & Sons, Inc, New York

Mankiw, N. Gregory . 2003. Teori Makroekonomi Edisi Kelima. Terjemahan. Jakarta: Erlangga Publisher.

Purwaningsih T, Erfiani, Djuraidah A . 2015. Comparison of Unifrom and Kernel Gaussian Weight Matrix in Generalized Spatial Panel Data Model. Open Journal of Statistics 5, 90-95.

Sambodo, Bayu Seto. 2014. Analisis Pengaruh Inflasi, BI Rate, Nilai Tukar Rupiah, dan Harga Emas Dunia terhadap Indeks Harga Saham Pertambangan di BEI (Periode 2008 – 2012): Jurnal Ilmiah Mahasiswa FEB UB Vol 2, No 2: Semester Genap2013/2014.

Sunyoto, D. 2007. Analisis Regresi dan Korelasi Bivariat. Yogyakarta: Amara Books.

TradingEconomics (2017). Inflation Rate, Export, Import Data. www.tradingeconomics.com. Accessed on January 16th 2018 at 16.45 pm.

Universitas Diponegoro. 2015. Analisis Spasial Pengaruh Tingkat Pengangguran Terhadap Kemiskinan di Indonesia. http://ejournal.undip.ac.id/index.php/media_statistika. Accessed on January 16th 2018 at 20.35 pm.

Wikipedia (2017). List of Countries by netural disaster risk. http://www.globalpetrolprices.com. Accessed on January 16th 2018 at 09.45 am.

Biographies

Tuti Purwaningsih is a Lecturer in Geo-Statistics and Big Data, Department of Statistics, Universitas Islam Indonesia. Ms. Tuti holds a Bachelor of Science degree in Statistics from Bogor Agricultural University and a Master of Science degree in Statistics from the same university. She is a Certified data Scientist with over 5 years of experience. Her passion in Data Analytics lead her to be a Statistics Consultant. She has various experience in local and international research project with Wahana Data Utama, ADB, World Bank, USAID IFACS, Findyr, etc. She is co-founder of Data Science Indonesia. She has taught courses in Geo-Statistics, Big Data, Geographic Information System, Business Environment, Regression Analysis, Project Management also Sampling Technique for her students and her community.